DOCUMENT RESUME

ED 140 645 FL 008 644

AUTHOR Culhane, P. T.
TITLE An Attempt to Classify and Evaluate Distractors in Tests of Listening and Reading Comprehension.
PUB DATE 76
NOTE 16p.; Paper presented at the annual conference of the British Association of Applied Linguistics (September 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Error Analysis (Language); Interference (Language Learning); *Item Analysis; *Language Instruction; *Language Tests; Listening Comprehension; Listening Tests; *Multiple Choice Tests; Objective Tests; Reading Comprehension; Russian; Second Language Learning; *Test Construction; Testing; *Test Items; Test Validity

ABSTRACT

Distractors, the incorrect responses to an item on a multiple-choice test, should be designed to create confusion in the minds of some students and to permit a competent student to be able to see that they are wrong. It is possible, by close scrutiny, to isolate the sources of this confusion and, by looking at a statistical analysis, to find out which types are most effective. This paper addresses itself to the first of these tasks. An attempt is made to classify distractors by finding out what they are dependent upon, and by inspecting the types of confusion they elicit. Examples are drawn from tests for students of Russian.
(Author/CFM)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED140645

An Attempt to Classify and Evaluate Distractors in Tests of Listening and Reading Comprehension

A paper presented at the Annual Conference of the British Association of Applied Linguistics (September 1976)

© P.T. Culhane

FL 008644

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
P.T. Culhane
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER

An Attempt to Classify and Evaluate Distractors in Tests of Listening and Reading Comprehension.

A multiple choice item typically consists of the following: a stem, which is sometimes a question and sometimes an incomplete sentence; several responses, one of which is the correct one. The usual number of choices is four although five is not infrequent. Tests with three choices also occur, and TRUE/FALSE tests. The latter are not popular because of the high importance of the guessing factor, i.e. random guesses would normally give 50%. An attempt is often made to circumvent this by adding a third choice CANNOT SAY. All the items I shall use as examples will be four choice items, with three distractors and one correct response in each item. They are taken from the 'O' level papers in Russian Reading and Listening Comprehension and I shall translate all the examples into English. The items used are all written by experienced item writers and have all been used in public examinations for the Cambridge Syndicate, the Joint Matriculation Board and other boards. We have statistics (provided by the Test Development and Research Unit in Cambridge) on each of the items and they are all satisfactory from the statistical point of view, but it is not my intention in this paper to concentrate on the statistical aspects, but on the nature of distractors and how they work.

One important thing about distractors is that they should work, by causing some students to choose them for reasons other than the correct one. The second important thing about them is that a competent student should be able to see that they are wrong.

The avoidance of mistakes in the writing of multiple choice items for public examinations is of paramount importance and there must be a careful vetting procedure before the items are finally used. The procedures adopted by the Cambridge Syndicate and later modified by the TDRU are carefully designed to avoid mistakes which can occur by aberration, inexperience or for any other reason. Items are written by a team and discussed by them before pretesting. After pretesting they are again sifted and items are modified or rejected according to their statistical performance. We are concerned at this stage with the FACILITY of an item and its ability to DISCRIMINATE between good students and bad ones. We also look at distractors to check that they discriminate in a negative way. If one has a POSITIVE distractor it means that the distractor has been chosen by more of the good students than by the bad ones, i.e. it does not contribute to the creation of a rank order. It therefore needs scrutiny, and it is not always possible to tell why the distractor has failed to operate correctly. Nevertheless, the distractor has failed and we usually re-write it to make it more evidently incorrect. The correct response may also be modified in order to change the statistical balance in its favour.

There are many ways in which distractors can be unsatisfactory. Many of the criteria used are applicable in examinations of any kind and concern such factors as appropriate level of language, length of papers, comparability between languages and so on. There are some ways, however, in which the nature of the faults is peculiar to this type of examination, that is, the distractor element of it. These are some of them:

1. A distractor may be correct if the writer has not seen that by a change of emphasis, intonation or

punctuation a meaning different from the one he had intended might be attributed to it. Punctuation is included here, as different languages tend to be more or less prescriptive in their rules. There are certain 'grey areas' in some languages, including English, in which for example commas may or may not occur according to the discretion of the writer. Russian tends to have commas where one would not expect them in English. The printed format of this type of test also adds to the complexity of the punctuation problem. I am not suggesting that punctuation should not sometimes be used as a distractor, merely that extreme caution should be exercised in using it.

2. A distractor may elicit a pattern of reasoning from intelligent students which make it a possible correct answer. An intelligent student might deduce that a distractor is correct if certain conditions prevail, even though these conditions are not stated in the passage. If it is likely or possible that these conditions exist, the item should be rejected. Students are however told to give the correct response 'according to the passage' and this is always taken into account when such distractors are discussed. The facility of the item is also important in this respect.

3. An item or distractor might reflect conditions which are unacceptably atypical of the L2 culture. To give an extreme example, an English Reader produced for Russian school beginners started with the words 'The Communist Party leads all children to a brighter future'. Whatever the pedagogical merit of such a sentence might be, it does not represent the sort of sentiment which is likely in an English environment. Another example from Russia at a more advanced level is the extraordinary sentence 'I wish to give my thanks to the head of the chair'. This was a sentence used in English by university students in concluding the defence of diploma dissertations. When the mistake was pointed out the Russians simply said that there was no English equivalent for what they wanted to say and this was the best one they could invent. In examinations of Russian as a foreign language we should avoid references to pubs, fish and chip shops and many more of our hallowed institutions.

Conversely, we should avoid items which rely too heavily upon a knowledge of the civilization element of the foreign country. It might be unfair to expect sixteen-year-old pupils to know that it is universally unacceptable for people to smoke in public transport, shops, cinemas and theatres in the USSR. It might be expected that they should know a little about the Pioneer organisation, or some other such elementary knowledge.

4. Another sense in which an item may be unacceptable is, as we said earlier, if the grammar or lexis is too difficult. Sometimes the content of the correct response is too difficult. The level may be lower in the distractors. The student might look at the distractors and find the nearest one to the correct response and choose that one. He may, on the other hand, reject all the distractors and choose the correct response without knowing what it means. If the boot is on the other foot and the correct response is very simple, the student might reject the correct response on the grounds that it is so simple that the cunning examiners could not possibly have written such a simple item. If this happens, the simplicity of the correct response is functioning as a distractor and the item should probably be rejected.

5. In some tests the correct response will relate to a whole sentence and the distractors to single words. I have seen examples of this in an EFL test, but not, I hasten to add, in any of our tests.

6. Some distractors or items might be unacceptable for social reasons e.g. they might remind a pupil that his relative has just died, fallen ill or met some other catastrophe. This sort of thing can conceivably slow a student down in an examination and thus mar his performance. Passages which evoke possible personal tragedies should therefore be avoided. 'Sick' items, that is, items which use 'sick' humour, should also be avoided because of their emotional impact.

7. One must also, at the first 'shredding' stage, try to exclude items which have a 'non-sequitur' as the correct response, and in passage items to exclude those which can be answered without reference to the passage.

A fairly standard way of writing distractors is to make an error analysis of the area being tested and give distractors which include the errors for that particular area. It is easy to predict what sort of mistakes an Arab will make in English and any of his mistakes are likely to be caused by interference of L1. The same would be true of any FL group of learners whose mother tongue is the same. Russians, for examples, always have difficulties with articles, as these do not exist in their mother tongue. English learners of Russian have difficulties with inflexions, verbal aspect and other specifically Russian features. This 'error analysis' approach is very valuable in spheres other than comprehension. In comprehension tests the writers' approach is usually intuitive and good items are produced by a 'semi-creative' act followed by rigorous selection. In Russian, L1 interference features do not predominate in comprehension tests. Most of the effective distractors work because of features either specific to the language being tested or because of features common to all languages. A comprehension test written in English and then translated would work as a comprehension test. This would of course exclude items which were dependent on L2 distractors and would not be an acceptable procedure, but nevertheless it would work. It is likely that the closer two languages are together, the greater the interference of L1 in comprehension tests. There will be more

cognate lexical items with different shades of meaning, the influence of one set of structures on the other will be greater, there will be more in common between the cultures. In investigating distractors in comprehension tests, I have not proceded from an error analysis, but I have turned the process on its head, perhaps with a view to investigating the problem from the other end (i.e. examining the various taxonomies of comprehension skills in relation to performance) at a later date. My present task is to take the process of item writing, look at the distractors, most of which are intuitively written and see what they seem to be relying on for their effectiveness. Having set myself this task I was provided with papers from TDRU and Cambridge, along with statistical analysis back to 1969. The papers (one each of Listening and Reading Comprehension containing 30 items each) appear twice a year. 120 items per year for eight years with three distractors to each item means an awful lot of distractors. I therefore resisted the temptation to examine each one, at least for the moment, and simply examined the tests until I was not obtaining any new features. Some of you may be familiar with the tests, others may not. I shall quickly outline them for you.

1. Reading Part I Vocabulary items, usually easy, followed by more complex continuations.
   Part II 3 passages with 5 questions to each.
2. Listening Part I Short items, again usually easy, with a request to the student to indicate either who is speaking, or where the person is, or what he is doing.
   Part II Short conversations with two questions after each one.

Situations. Each situation is followed by five questions relating to that situation.

Part III Dialogue + 4 or 5 questions

Narrative + 4 or 5 questions

The whole of the listening test is recorded with appropriate spaces for the student to mark his response. Experiments have been done in Spanish with 'all-heard' items. In other languages the responses are printed on an answer sheet and the student is asked to indicate the correct one.

For the purposes of this paper, I have concentrated mainly upon items in Reading Comprehension rather than Listening. The same sorts of factors operate in Listening as in Reading, but the level of language tends to be lower in listening tests. Additional factors such as aural discrimination operate and this causes us to make the distractors rather less complex.

I want to discuss distractors on 3 parameters. The first of these is relatively straightforward (labelled 'A') and might be called DEPENDENCE. A distractor is dependent on one or more of the following:-

A1. A passage

A2. The stem of the item.

A3. The correct response.

A4. Another distractor

I have seen some tests (particularly EFL tests) in which a student is asked to infer or deduce information relying both on the passage and upon information not presented in the passage. In these cases I would place the distractor in the 'passage-dependence' category and perhaps suggest a sub-division of that

category, involving assumed knowledge. This sort of item is fairly common in, for example, placement tests for science students in foreign universities where a knowledge of English is essential and what one wants to know is whether a student can handle knowledge of his scientific subject in the L2.

Having decided that there are four basic categories of dependence it was then a simple task to look through each item and mark each response with the appropriate number. Some distractors had more than one dependence in that they could be both passage and stem dependent, but usually the dependence did not extend beyond two.

I then had a look at the types of CONFUSION which were likely in the students' minds and which appeared to be intended by the author. A student, in opting for the wrong answer does one or more of the following:

a) He ignores the correct response, either because he does not understand it or for some other reason.

b) He chooses an answer because he thinks it is more likely to be correct than is the correct response.

c) He makes a random guess.

Distractors, as we have said, should have two characteristics:

1. They should be wrong, and quite clearly so, to someone who has understood the item.
2. They should work, in that they should have in them an element which causes students who do not fully understand the passage to choose them.

The second and third parameters tend to run in parallel in that they have roughly the same characteristics. The two categories are classified under the heading of CONFUSION. Confusion implies two things:

a) that a student has chosen a distractor for some positive reason, and

b) that the distractor is wrong for some 'negative' reason.

To take a simple example a student may choose a distractor because it is similar in form to the correct response or to somethingning in the stem or passage. But his choice is wrong because of some temporal, personal or other confusion, or because the student had looked at the wrong part of the passage. There is a parallel here with crime or sin. If someone robs someone or commits some misdemeanour he usually does it for some positive reason, i.e. he wants to be rich or gain pleasure. He is wrong in doing so, because he makes someone poor, causes suffering or for some other negative reason. The positive aspect I have labelled B and the negative C. They are parallel in the sense that they appear to subdivide into the same sorts of categories. They are not parallel in the sense that if the choice is made for, say, 2 reasons the negative is likely to appear somewhere else in the list C5,6, or any other 'C' number.

I shall give the categories of confusion in fairly random order, and shall try to indicate their importance subjectively. So far I have not tried to assess the importance of the categories statistically, and I am not sure how far I should do so, as I

think that it might be wise simply to use them to try and find out more about comprehension skills in general. Bear in mind that there is a positive and a negative side to each of the categories each of which is a part of a separate parameter. So far I have isolated 12 types of confusion, but there may be more. Just as a distractor might be dependent on one or more part of the test, so it may cause more than one type of confusion. Some of them have three or four.

B1 Has the general label 'semantic' and, if you attach the stem to the distractor you are left with a non-sequitur.
(In these examples I shall give the stem and the distractor, without the correct response.)
e.g. A Muscovite is standing on a Moscow railway station.
He asks a porter what time the show begins.
The ground there yields a rich harvest because the Kolkhoz workers have a lot of money.

B2 Form. This type of confusion usually exists when distractors are used which have a similar form to the correct response or a word associated with it. These are common in vocabulary items. The similarity may be homonymic, but not necessarily fully so. A fairly common one in Russian vocabulary is with the words for 'to bathe' and 'to buy'. When a bathing situation is evoked, whether or not the actual word for 'to bathe' is used, the buying element is very often included as a distractor.

B3 Contra. This is a very important and very common category where the distractor used is either a complete opposite, a negation of the correct response or a partial negation. There is a cline here.
A simple illustration of this is
'Let's open the window, the room is very cold.'

Or:

It's summer, it is warm. The time is two o'clock in the afternoon. Misha is sitting by the window and reading a book. His mother says, 'You must be cold sitting by that window.'

The above two examples could also be examples of B1 and could have a vocabulary element. Perhaps a more appropriate illustration would be the following items associated with a passage.

In this example a sentence in the passage tells us that in a certain area in the Soviet Union there is a good deal of coal, oil, iron and some copper. The distractor says that the area is rich only in oil and coal. In another passage we are told that a man had to go without breakfast, the distractor tells us that he had breakfast quickly. The reason the distractor works is because of a similarity of form and the confusion has occurred because the student has not noticed the 'contra' element.

B4 Temporal. Not necessarily tense confusion. 'Come and see us tomorrow. Let's say two hours ago.'

This often reflects on the sequence of events in a passage where distractors operate for some other reason - form, because the student realises that they refer to time but does not have the vocabulary to discriminate between correct response and distractor, or for any other reason.

Straight tense confusion.

Tanya is preparing lunch in the kitchen. She says:
'It's already eight o'clock and Vitya has not got up yet.'

B5 Personal. When the student chooses the distractor because he has chosen the wrong person taking an action or making a statement in the text.

E.g. A student is looking for a book in a library. He asks the librarian "What subject do you want to study here?"

B6 Grammatical. In Russian this very often involves such things as verbs of motion and other problem areas for non-slavonic speakers. This is an area where L1 interference can have a significant part to play.

B7 Vocabulary. Again L1 factors come into play, but do not predominate since distractors are usually dependent on some similarity of 'form' to correct responses.

B8 Deductive. Possible but not indicated. Usually passage dependent items. This sort of distractor might attract the stronger students and must be contrasted with a fairly strong correct response if it is to work properly.

e.g. A great scientist has been working for many days in his laboratory. His wife says to him, "Why don't you work in bed?' Any ambiguity is caused by L1 interference!

B9 L1 Interference - Common in vocabulary items where there is a part-cognate.

B10 Contextual. Attraction to the wrong part of the passage, e.g. when the distractor contains an element similar in form to something in the wrong part of the passage to that about which the question is being asked. This type is very common in passage items.

B11 Place confusion, e.g. Moscow may be mentioned as the capital of the Soviet Union and another town as the main steel centre. The distractor will say that Moscow is the main steel centre. More sophisticated examples are frequent. 'When John told Mary it was all over she ran into the kitchen.' when in fact we know from the passage that she was in the kitchen to start with and ran out of it.

B12 Punctuation. Items which rely on punctuation should have distractors which are quite clearly distinguished by punctuation and which make it clear to the student that the problem is one of punctuation e.g. by having very similar form.

As I have said, this work is by no means complete. I have found the study fascinating from many points of view and have given myself greater insight into examining distractors from a practical point of view. I realize I have not yet fully isolated the 'C' parameter which has 'ghost' qualities. I would be most grateful if the discussion were to help me in the furtherance of my research and suggest where this might be extended although disappointed if, to use David Chrystal's words, you were to say 'You won't get anywhere like that'.